

Storage Policies for the Illinois Campus Cluster

Policy Author: Jonathan Greenberg (jgrn@illinois.edu)

Passed: 12/18/2013 (4 to 1)

Overview and Justification: as the Campus Cluster is a shared resource, storage is paid for by individual investors. Currently there are no explicit rules for dealing with storage usage (either at a User level or an Investor level). This policy is to explicitly define these policies, and provide courses of action in the case that an individual user or an investment group exceeds their purchased storage quota.

Technical Details:

Definitions:

Illinois Campus Cluster Operators: The team responsible for installing, operating, and maintaining Illinois Campus Cluster Instances and for providing basic support to Users.

Investor Group (“Investors”): Faculty and/or units who, individually or as a group, invest in shared computing resources as part of the Illinois Campus Cluster Program. As referred to in this document, “Investor” is assumed to include both the Investor and any designated representatives. The Investor, or their designee, represents their interests as part of the Investor Forum.

Users: Faculty, staff, students, and collaborators who have access to the computing services of the Illinois Campus Cluster Program. A User must belong to at least one Investor group and at least one Investor project.

Storage Allocation Definitions:

Investor Total Storage (“ITS”): the total amount of usable storage purchased by a single investor group, which may include both disk purchases as well as the per-node storage allotment.

Investor Home Storage (“IHS”): the amount of the ITS that is allocated to Users’ home directories, which are governed under the “Policy on Appropriate Use of Computers and Network Systems at the University of Illinois at Urbana-Champaign” (<http://www.cam.illinois.edu/viii/VIII-1.1.htm>).

Investor Project Storage (“IPS”): the amount of the ITS that is allocated to locations other than User home directories (e.g. /projects/[investor group], /usr/apps/[investor group]).

Note: ITS = IHS + IPS

Investor-level storage policies

At all times, an Investor must be guaranteed 100% use of their purchased space.

Investor-level total storage usage (“ITSU”) is defined as the Investor Project Storage Usage (“IPSU”) plus IHS. Note that actual usage of the Home Storage locations are not considered, only the allocated amount. An investor is expected to keep their usage (ITSU) under their investment (ITS). If ITSU exceeds ITS, this may trigger a hard quota event that will prevent further use of the disk space. At 95% ITSU, an Investor and their technical representative will be notified via email that they are approaching their quota. If ITSU is between 100% and 110% of ITS, this will trigger a soft quota with a 14 day grace period during which the Investor must bring their ITSU below ITS, or purchase additional disk space. Once the grace period expires, or if ITSU is greater than 110% of ITS, a hard quota limit will be enacted and no further writes to the Investor’s home or project space will be permitted until ITSU is reduced to below ITS or additional storage is purchased. The Operators reserve the right to modify the upper end of the soft quota amount (by default, 110%), should it become necessary for the continuing operation of the Campus Cluster and to meet the obligations of providing 100% of an Investor’s purchased space at all times.

It is the Investors’ responsibility, not the Campus Cluster Operators or the Executive Committee, to make sure their users are keeping their storage usage within the limits of their investment and, when needed, educating and encouraging their users on proper usage of the purchased storage space and/or purchasing additional storage. The Illinois Campus Cluster Operators can provide tools to help identify and manage Users’ storage usage, but the Investor is responsible for dealing with this proactively.

Project-level storage policies

By default, an investor will have a single shared project space (/projects/[investor group]/shared) which will be referred to as the Investor Primary Project Storage (IPPS) which will initially have an allocation equal to the total IPS. An investor may, by request, allocate some of their IPS to other sub-projects (Investor Sub Project Storage, “ISPS”). The sum of the ISPS allocations plus IPPS may not exceed the total IPS. Note that home directory allocations (IHS) are subtracted from the IPPS allocation. Each sub-project can have its own quota and soft/hard-lock procedure.

User-level storage policies

A User has three basic storage pools they may use on the Campus Cluster: user home directory storage (“UHS”, a fraction of the IHS), project space (IPS), and scratch space. Each storage pool has different policies/regulations as follows:

User Home Storage: a User, by default, will be allocated 2GB of UHS. Total investor home storage (IHS), then, is the sum of all Users’ UHS within a given Investor Group. As such, each new User added to an investment group will reduce the available project storage (IPS, or IPPS if subprojects are present) by their home directory allocation. An investor may request from the Illinois Campus Cluster Operators a higher UHS on a User-by-User basis or for the entire Investor User group, but must recognize that unused space by a User is still subtracted from the

available space to the Projects. Thus, an Investor should generally not request an increased UHS and instead encourage Users to take advantage of the Project and Scratch spaces.

Once a Users' UHS is exceeded, a hard quota will go into effect, and a User will no longer be able to write to their home directory until the UHS usage is reduced to under the UHS.

Access to the UHS is governed by the campus Appropriate Use Policy (<http://www.cam.illinois.edu/viii/VIII-1.1.htm>). This policy restricts an Investor's ability to access or delete data in Users' UHS without the Users' permission, even after a User has left the University. This limit is one of the chief reasons why an Investor should be cautious in allowing the UHS quota to be raised and should instead encourage the use of IPS space. This restriction does not apply to IPS space. An investor and the operators can only archive UHS and move it to a secure location, they cannot examine or delete the data without authorization from the User or from campus. To access or delete data without a Users' explicit permission, whether it is in the UHS, an archive, or an alternate secure location, an Investor must follow the procedure documented in section VI.C of the Appropriate Use Policy:

- <https://wiki.cites.illinois.edu/wiki/display/SecurityKnowledge/Access+to+Employee+Email+and+Other+Files>

User Project Storage: By default, a User will have access to a shared project directory (/projects/[investor queue]/shared/[UID]). This directory will be read-only for all other Users in the Investor Group, so should not be considered private. Data in this and other project directories may be accessed, archived, or deleted at the Investor's discretion. Total investor project storage usage (IPSU) is the sum of the individual user project storage usage (UPSU). The per-User Project Storage quota (UPS) policy, by default, will be set to prevent a single user from causing a soft or hard quota lock on the IPS, and will be calculated as: $75\% * ITS / (\text{number of users})$ where "number of users" is the current or projected number of users, defined by the investor (thus, it is not dynamic). This default retains 25% of the ITS to be reserved for adding additional Users or to increase individual Users' quotas on a case-by-case basis. While an Investor may request other quota strategies, UPS may never exceed IPS. We provide two additional easy-to-implement strategies that an Investor may choose:

- No UPS quota: $UPS=IPS$. Each User has access to all of an Investors' IPS thus User Project Storage (UPS). This means that any one User can fill up the entire IPS thus triggering a soft quota situation. This was the default setting under the previous CC quota system (up to Fall 2013).
- Liberal UPS quota: $UPS = 110\% * ITS / (\text{number of users})$. Users can theoretically initiate a soft quota situation, but not a hard quota situation.

The Investor, at any time, shall retain the rights to move, delete, or modify (e.g. archive) data within their IPS. Any agreements or issues about the legal ownership of data falling within an Investor's IPS are the responsibility of the Investor, and the Investor alone.

Scratch Storage:

All users will receive personal space in scratch at the path '/scratch/users/<netid>'. All jobs will also receive space in scratch at the path '/scratch/jobs/<jobid>'. At a '2 to 1' ratio of time, preference is given to job folders over user folders. Job folders will be kept for a guaranteed two weeks, and up to four weeks dependent on available space, based on maintaining at least 20% free space on the scratch filesystem. If there is a need to go shorter than two weeks, a day's notice will be given to users before the scratch space is purged. Jobs folders are cleaned up based on end time of job, not the age of the files. Files in the users folder are cleaned up based on age. Folders are not deleted from user's folders but all files and folders in a job folder are removed. Exceptions can be granted by request to the Operators.

Temporary Grant of Space: There may be situations in which an Investor or User may need additional temporary, but non-scratch, storage space for a particular task. In this case, the Investor can make a request to the Operators for a temporary increase in their ITS beyond their purchased amount. The request should include 1) How much additional storage is needed, 2) a justification for the increase, 3) how long the increase will be needed for, and 4) a clear description of how the space will be freed up at the end of the grant period. This request cannot exceed 10% of total available free space on the CC, and/or double the Investor's ITS, whichever is smaller, and the increase can be for no longer than 1 month per year. Additionally, the space cannot be used for IHS (home directory) increases, only IPS increases. The Grant of Space is under the Operators' discretion and will be determined on a first-come-first-serve basis, as is revoking the Grant of Space before the requested period is over should it be deemed necessary. If the Grant is approved, the Investor ITS will be increased by the requested amount. If, at the end of the grant period, the Investor has not reduced their usage to under their purchased ITS, a hard quota lock will immediately go into effect until the storage is ITSU is lower than ITS.

Tools for Investors

The operators will be tasked with providing tools for investors to establish and manage their quotas. All questions about these tools should be directed towards the Operators.

Proposed Timeline: announcement of policy December 14, 2013 and emails sent to all investors over their quota that they have 2 months to reduce their usage or purchase more storage. January 14, 2014 for CC Operators to finish testing the implementation of the policy. February 14, 2014 for full implementation of the policy.

Impacts of the Policy:

Users: after the policy is implemented, Users may find themselves unable to write to their home and/or project spaces. Additionally, Users will need to adjust to a project-based file permissions, in which other users in the same project may have read/write access to their data.

Investors: after the policy is implemented, Investors who continue to exceed their purchased storage quota will be at risk of a hard-lock, in which none of their Users will continue to be able to write to either their home directories or their project directories. This policy will also put

additional responsibilities on the Investor to monitor their disk usage and work with the Operators to set up permissions on their project spaces.

Operators: there will need to be potentially significant pre-policy work setting up quota rules and new permissions. Investors and users that are locked (over quota) may go to help@campuscluster.illinois.edu first to inquire about the lock-down. Operators will need to enforce the policy that proper use of the storage is the Investor's responsibility, not the Operators or Executive Committee.

Estimated Costs to Implement the Policy: additional hours from the Operators to set up the quota/permissions system and additional tools.